THE EFFECT OF MONETARY INCENTIVES ON MONITORING PERFORMANCE

Kenneth Lee Yufer

United States Naval Postgraduate School

THESIS

THE EFFECT OF MONETARY INCENTIVES
ON MONITORING PERFORMANCE

by

Kenneth Lee Yufer

October 1969

The Effect of Monetary Incentives on Monitoring Performance

by

Kenneth Lee Yufer
Lieutenant, United States Navy
B.S., United States Naval Academy, 1962

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
October 1969

ABSTRACT

A vigilance experiment was performed to evaluate the effect of monetary incentives on visual monitoring performance. The vigilance task was the detection of a slightly larger excursion of a voltmeter needle making 50 uniform excursions per minute. The length of the vigil was 48 minutes, during which 32 signals were presented. Ten subjects (*S*s) in a control group performed the task without possibility of reward. A second group of 10 *S*s performed the identical task receiving monetary rewards based on performance. Subjects receiving monetary rewards (M) detected significantly more signals ($p < .001$) than did the control group (NM). A significant time decrement ($p < .001$) and a significant interaction between group and time existed ($p < .001$). The detectability (d') or sensitivity for the signals remained essentially invariant for both groups, although slightly greater for group M, and the criterion level (beta) was comparatively lower for group M. The cost factors were effective in manipulating monitoring performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION AND BACKGROUND

Extensive research has been conducted concerning vigilance tasks in which man is required to detect randomly and infrequently occurring events. The vast majority of this effort has been concerned with discovering what physical-task parameters may be responsible for vigilance decrement, i.e. the decrease in vigilance performance. Recently, however, some attention has been devoted to motivation as the key to understanding vigilance behavior. A catalyst precipitating this line of reasoning was the fact that performance improved with "knowledge of results", i.e. telling subjects (Ss) when correct detections and false alarms were made. (Wiener 1963, and others). In actual monitored systems, however, knowledge of results would be unavailable, or at best extremely delayed. If the feedback circuitry, human or electronic, were capable of providing immediate knowledge of results, the occurrence of a signal would have to be known, and the need for a human monitor would be eliminated.

Several other methods of motivation have been attempted with varying degrees of success. Baker, Kowal and Ware (1964) pointed out that experimenter (E) attitude toward the Ss significantly affected their performance on a vigilance task. Several studies (Pollack and Knaff, 1958; Sipowica, Ware and Baker, 1962; Bergum and Lehr, 1964; Levine, 1966; and Smith, Lucaccini and Epstein, 1967) have employed monetary incentives to increase motivation.

Pollack and Knaff (1958) reported a slight improvement in detection performance when Ss were given the incentive of "an extra hour of pay" for either 100 percent detection in seven of eight 10-minute periods or the greatest improvement in scores over previous tests. Subjects were guaranteed pay for 50 hours of service and the incentive of "an extra hour of pay" was not quantified.

Sipowicz et al. (1962); Bergum and Lehr (1964); Levine (1966); and Smith et al. (1967) all provided rewards in proportion to the number of targets detected, minus varying punishments for missed signals and false alarms. All experiments yielded some degree of improvement in results as compared to a non-motivated group.

These monetary incentives were, however, subjectively arrived at. Classifications such as "low incentive", "medium incentive", and "high incentive" were used in a subjective manner. Utility and/or decision theory may be used to help quantify the experimental incentives used.

Utility theory suggests that the costs and values of deciding whether or not a signal is present can be illustrated by a utility decision matrix:

UTILITY DECISION MATRIX

	S	$\bar{S}$	EV
	π_1	π_0	
R	+	-	A
$\bar{R}$	-	+	B

S = signal present $\bar{S}$ = no signal present

R = response $\bar{R}$ = no response

π_0 = a priori probability of no signal being present

π_1 = a priori probability of a signal being present

EV = expected value (average utility) of decision

The +'s and -'s represent the utility points associated with a particular action. The S theoretically would seek to maximize his EV and therefore one should assign incentives such that A is positive and A > B. If, for example, A = -2 and B = 3, S would decide to "not respond". However, Galanter (1962) found that at best a logrithmic relationship existed between money and utility, making the assignment of utility points very difficult.

A second approach is to use the statistical tool of decision theory, and by applying one of the developed criteria for an optimum decision rule, arrive at the proper costs to apply in order to minimize the risk of making a decision.

Direct application of the above two approaches would be very difficult however, since, as shown by Egan, Greenberg and Schulman (1961); Weiner, Poock and Steel (1964) as re-analyzed by Taylor (1965); Levine (1966); and others, the vigilance decrement is not due to a change in the efficiency of the S's senses but rather a shift in the S's criterion for response in the direction of greater conservatism. In other words, the detectability of the signal remains constant, but it is the S's criterion that changes. This means that the S's utility decision matrix or the costs and/or the probabilities of a signal being present, change with time. Thus the S's decision rule changes as a function of time. Deese's expectancy theory (Deese 1955) states, in brief, that the likelihood that the subject will respond to a randomly occurring signal depends upon his expectancy about the appearance of a signal. This expectancy is built up as a kind of "averaging process" based on past signals, implying that to the observer, the probabilities of signal presence change with time. Several studies (Jerison, Pickett and Stenson, 1965; and Mackworth, 1968) found that an increased signal rate did actually decrease the probability of detection.

This study will assume fixed a priori probabilities in order to isolate the effect of monetary costs on motivation, and ultimately vigilance performance. However, additional research in support of Deese's theory certainly seems warranted.

The current study will attempt to decrease the vigilance decrement by inducing motivation through statistical assignment of the costs of the decision parameters in order to optimize the observer's decision criterion. These costs of the decision parameters will be constant, although further study should be made to determine the significance of this type of value assignment.

In order to manipulate the costs of making a decision as a function of time in a laboratory environment, a procedure similar to knowledge of results would have to be employed. Therefore this study will assign fixed costs to the decision parameters. Although internal inhibition or caution apparently increases with time, when a S̲ does become more reluctant to respond to the presence of a signal his non-optimal payoff should dictate that he is no longer maximizing the expected value of his decision rule and the inhibition should dissipate. Thus, the S should return to the original decision costs.

Signal detection theory (Green and Swets, 1966) was applied in the current experiment. This theory is based on the concept that signals are detected against a background of "noise". This noise represents random variations in the channel through which the signal reaches the decision function. Any specific observation may come from one of two populations, either noise alone or noise plus a signal, and the observer must decide whether or not he will accept that observation as a signal. The difference between the means of the distributions of the two populations, noise and signal plus noise, is the detectability of the signal. When the two distributions are Gaussian and of equal variance, the difference between the means divided by the standard deviation is designated d'. The measure of the sensitivity of the observer for the signal is independent of his decision criterion. The decision criterion is measured by the parameter β, where the acceptance region for a positive response consists of all events whose likelihood ratios are equal to or greater than β. An increase in β indicates an increase in caution by the observer.

It will be assumed that the detectability of the signals, d', is essentially constant as verified for the Jump Clock Test by Mackworth (1968) and in general by Loeb and Benfold (1964) and Levine (1966).

The costs for the decision parameters were calculated using the Neyman-Pearson decision rule which yields a specific likelihood ratio value, β, for a specified a priori probability of false alarms (see Selin, 1965, Chapter 2). This specific value of β was then used in the Bayesian decision criterion in order to obtain a ratio of the costs for the decision parameters. The Bayes criterion consists of minimizing average cost when the a priori probabilities are known and the average cost is a linear function of the absolute error probabilities (see Selin, 1965, Chapter 2). By assigning a specific value to the probability of false alarms ($\alpha = .05$) ultimately a ratio of costs can be determined. In other words the decision maker is maximizing the acceptance of a signal when there is a signal present, while holding a constant probability of false alarms.

The actual calculations are as follows:

$f(x|s)$ = p.d.f. of signal plus noise

$f(x|n)$ = p.d.f. of noise alone

C_{00} = cost of correct rejection

C_{01} = cost of miss

C_{11} = cost of correct detection

C_{10} = cost of false alarm

π_0 = a priori probability of no signal being present

π_1 = a priori probability of a signal being present

α = probability of false alarm

β = critical value of likelihood ratio

Assume $f(x|n)$ is $N(0,1)$ and $f(x|s)$ is $N(\mu s,1)$; then for this case, by signal detection theory:

$$\alpha = \int_{\beta}^{\infty} f(x|n)\,dx$$

This experiment will assume $\alpha = .05$ $\beta = 1.645$ (from table of normal distributions).

Using the Bayes decision criterion likelihood ratio

$$\frac{f(x|s)}{f(x|n)} \geq \beta$$

$$\beta = \frac{\pi_0 (C_{10}-C_{00})}{\pi_1 (C_{01}-C_{11})}$$

π_0 and π_1 were determined from the experimental target presentation, thus the known values were:

$$\pi_0 = .9867$$

$$\pi_1 = .0133$$

$$\beta = 1.645$$

rearranging the above known values yields

$$.022 = \frac{(C_{10}-C_{00})}{(C_{01}-C_{11})} \quad .$$

The theory thus yields an infinite number of possible combinations that satisfy the above ratio. For this experiment, the following combination of values was chosen:

$$C_{10} = 2$$

$$C_{00} = 3$$

$$C_{01} = 4.55$$

$$C_{11} = 50 \quad .$$

Any assignment of specific costs to C_{00}, C_{01}, C_{10}, C_{11} implicitly indicates some form of utility assessment. It is recommended that a further study be conducted using various ranges of costs with a fixed ratio, β, to determine if there is any significant difference in performance as a result of the variation of the assigned costs of the decision parameters.

As stated, human behavioral theory sometimes suggests that subjective probability, not a priori probability, and utility, not monetary value, determine the S's decision. If we assume nonlinear transformations of the values of the decision parameters and of the a priori probabilities, the maximization of the subjective value would still have the form illustrated above. The only thing that would change would be the value of β. The likelihood ratio criterion would still be the optimal decision rule (Green and Swets, 1966).

II. METHOD

A. APPARATUS

The monitoring task consisted of the detection of an abnormally large deflection of a voltmeter needle which made 50 deflections per minute. The normal deflection and the signal were produced by electrically energizing the meter. The signals were programmed on a paper tape and stepped through an Ohr-tronics 8-channel tape reader at a rate of 50 characters per minute. Each character produced either the short background deflection of 25-degrees, or a signal, a 32-degree deflection. The impulses appeared on a meter consisting of a uniform white background and a black needle. The impulse circuit and damping characteristic of the meters were such that the impulse appeared as a rapid rise and fall of the needle with no pause or "bounce" at the peak, or bottom. The needle returned to its resting position before the next deflection. The S responded by pressing a silent hand switch which activated a pen on a Lafayette Model 5040 multi-pen recorder. Signals were also automatically recorded in the same manner. Subjects wore earphones which played white noise at about 60 db to further assure auditory isolation.

A schedule of 32 signals for the 48-minute run was determined from a table of uniformly distributed random numbers. The only restrictions on randomness were that the minimum inter-signal interval was 0.3 minutes and that 8 signals would appear on each 12 minute block.

B. SUBJECTS

The Ss were 20 military officer graduate students at the Naval Postgraduate School. None had served previously in monitoring studies.

C. PROCEDURE

The Ss were randomly divided into Group NM (non-motivated) and Group M (motivated), consisting of 10 Ss each. Identical instructions were read to each S at the beginning of the experiment explaining the nature of the task and the response procedure. In addition, Group M was given an explanation of the cost incentive matrix. This explanation included the fact that it was possible to earn between 10 and 15 dollars for a perfect performance. A maximum effort actually yielded $12.72, a conversion of .75 cents per cost matrix point. However, to insure against any knowledge of signal presentation, the previously mentioned dollar range was given. A practice session consisting of 9 signals in rapid order followed in which knowledge of results was given. The E warned the Ss that the signals would not appear so rapidly in the actual experiment. The S was asked to remove his watch and the task was then begun with no interruptions permitted.

III. RESULTS

A response within 2.0 seconds after presentation of a signal was scored as a detection; all other responses were scored as false alarms. Figure 1 shows the mean percentage of signals detected as a function of time periods for both groups. These raw detection percentages were transformed to radians by the arcsine transformation in order to satisfy the normal distribution requirement for analysis of variance. A nested-factorial analysis of variance was performed on the radians, with Ss nested into groups but common to the four time periods, see Table I. The difference between groups was significant, $F(1,18) = 16.43$, $p < .001$, with the motivated group detecting a higher mean percentage of signals, Group M 86.9%, Group NM 68.5%. The analysis showed a significant decrement over time periods, $F(3,54) = 7.80$, $p < .001$, as shown in Figure 1. A significant interaction existed between groups and time periods, $F(3,54) = 8.50$, $p < .001$, due to a significantly higher percentage of detections by the motivated group than the non-motivated group.

Figure 2 displays the mean percentage of false alarms as a function of time periods for both groups. The percentage of false alarms was calculated utilizing the fact that it was possible to have 592 false alarms per 12-minute block, i.e. 600 deflections minus 8 signals. A non-parametric median test, designed to measure whether the two independent groups differ in central tendency, indicated there was no significant

TABLE I

ANALYSIS OF VARIANCE OF ARCSINE TRANSFORMATION OF SIGNALS DETECTED

Source	df	ms	F	p
Between Ss	19			
Groups (g)	1	5.6196	16.434	<.001
Error (between)	18	0.3420		
Within Ss	60			
Time Periods (T)	3	1.5362	7.795	<.001
G x T	3	1.6751	8.500	<.001
Error (within)	54	0.1971		
Total	79			

difference between the means of the two groups at the 0.001 level, $\underline{\chi}^2(1) = .051$, $\underline{p} < .001$. The experimental range of percentage of false alarms, 3.6 to 1.0, approximated the a priori percentage of false alarms, .05.

The calculation of d', the statistical difference or distance between the means of the signal and signal plus noise distributions of equal variance, was obtained from a table compiled by Elliot (Swets, 1964), using mean percentages of signals detected and of false alarms per group per time period as suggested by Mackworth (1968).

The parameter β was calculated by evaluating the likelihood ratio at the points indicated by the mean percentages of signals detected and of false alarms, i.e. evaluate the ratio of the normal probability density functions $f(x|s)/f(x|n)$.

The experimental values of the d' and β parameters for both groups are shown in Figure 3. These results indicate that d' remained essentially invariant for both groups, being slightly greater for group M; while except for time period 1 where β was approximately equal for both groups, β increased more rapidly with the NM group. This indicates that the detectability of the signals was approximately equal for both groups but, after the first time period the NM group was comparatively more reluctant to respond to a signal.

Taylor (1965) indicated that β itself does not measure the placement of the criterion on the scaled log-likelihood-ratio axis, but rather the parameter (log β)/d' should be evaluated. The decision need not be based on likelihood ratio per se, but the decision rule must be equivalent to a likelihood-ratio criterion in order for the relationship, $d\,P(x|s)/d\,P(x|n) = -f(x|s)/-f(x|n) =$ likelihood ratio, to exist (Green and Swets, 1966). A plot of (log β)/d', Figure 4, indicates the same trend as β in Figure 3.

IV. DISCUSSION

The results of this study provide support for the hypothesis that when motivation is increased by monetary incentives, performance in a simple vigilance task can be significantly enhanced. This is pointed out by the fact that the mean percentage of signals detected, 86.9%, by the motivated group (M) was significantly higher than the mean percentage of signals detected, 68.5%, by the non-motivated group (NM). Similar results were reported by Pollack and Knaff (1958); Sipowicz et al. (1962); Bergum and Lehr (1964); Levine (1966); and Smith et al. (1967).

The hypothesis that the detectability of the signal, d', is essentially constant was supported. For the first time period, d' was equal for both groups. During the remaining three time periods a slight overall rise in d' was found for the motivated group and a slight decrement for the non-motivated group.

The decision criterion, as measured by beta, showed an increase in caution for both groups during each time block. During the first time block, beta was essentially the same for both groups. However, for the three remaining time blocks, beta increased more rapidly for the non-motivated group. This is consistent with the fact that the motivated group performed significantly better than the non-motivated group.

The signal detection theory analysis of the results is consistent with those resulting from the analysis of variance.

Signal detection theory states that the significant difference in performance between the two groups resulted from the fact that the non-motivated group was more reluctant to respond to a signal when one occurred, while the detectability of the signals was essentially the same for both groups.

It is recommended that in future studies utilizing monetary incentives a statistical evaluation of decision costs be conducted using decision theory and a Bayesian type decision criterion. The statistical procedure used in this study not only provides a statistically determined cost ratio but also allows the E to choose an a priori probability of false alarms.

In summary, the performance in a simple vigilance task was significantly enhanced when motivation was increased by statistically calculated monetary incentives. The theory that the detectability of the signal remains essentially constant and that the increase in caution with time is responsible for the vigilance decrement was supported. It is believed that a theoretical model of vigilance behavior must include motivational parameters.

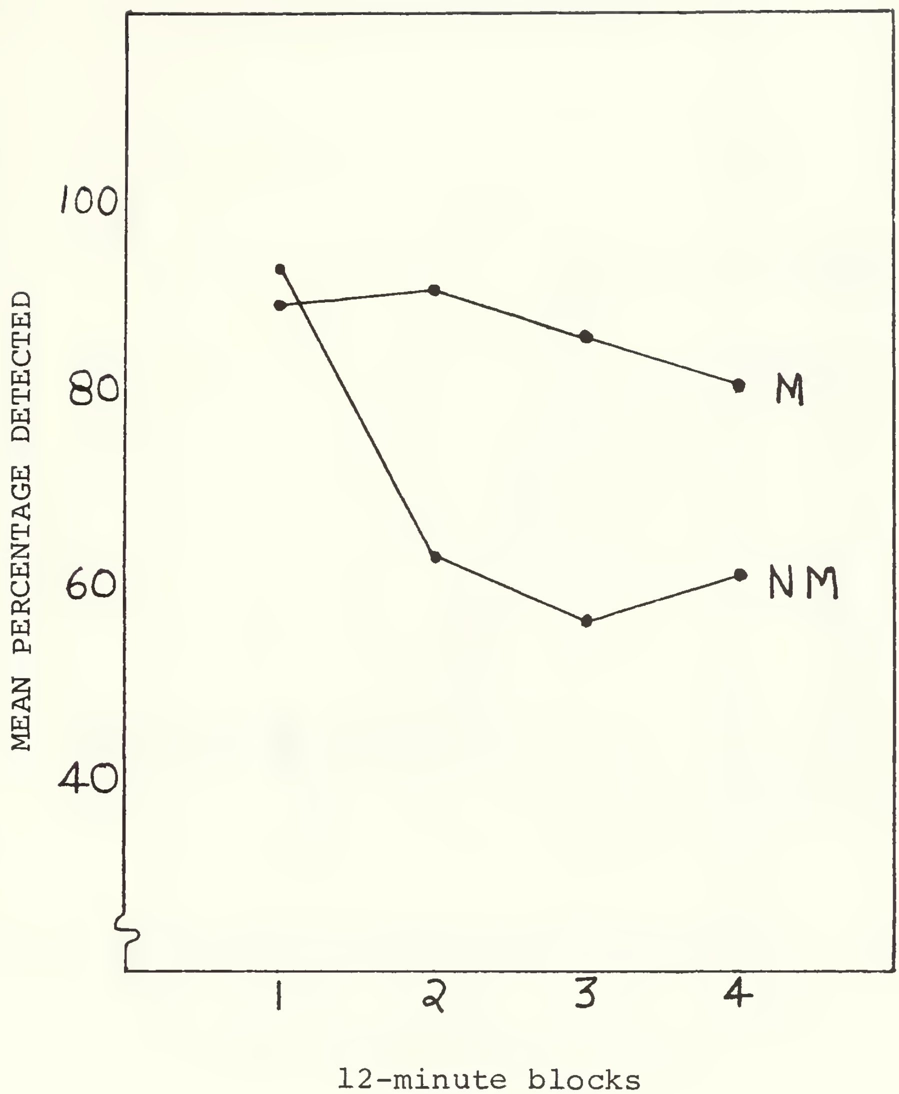

FIGURE 1: MEAN PERCENTAGE OF SIGNALS DETECTED AS A FUNCTION OF TIME

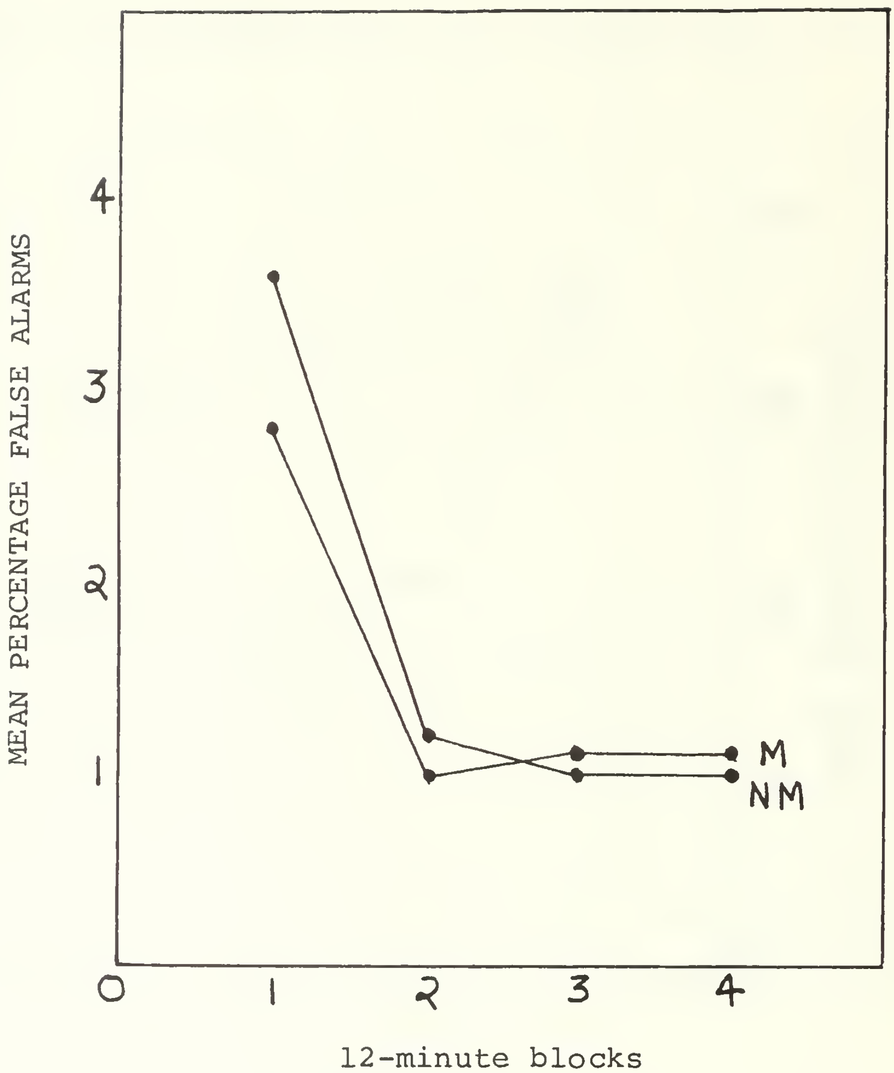

FIGURE 2: MEAN PERCENTAGE OF FALSE ALARMS AS A FUNCTION OF TIME

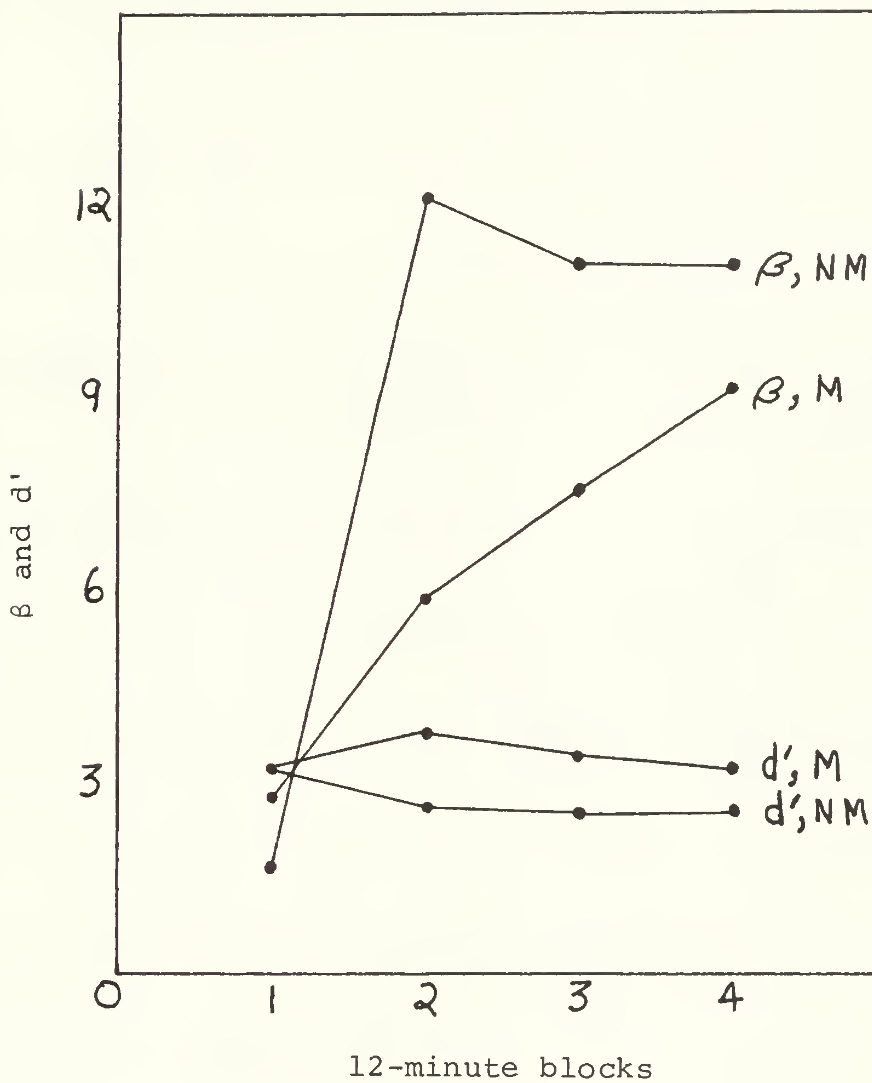

FIGURE 3: THE QUANTITIES d' AND β
AS A FUNCTION OF TIME

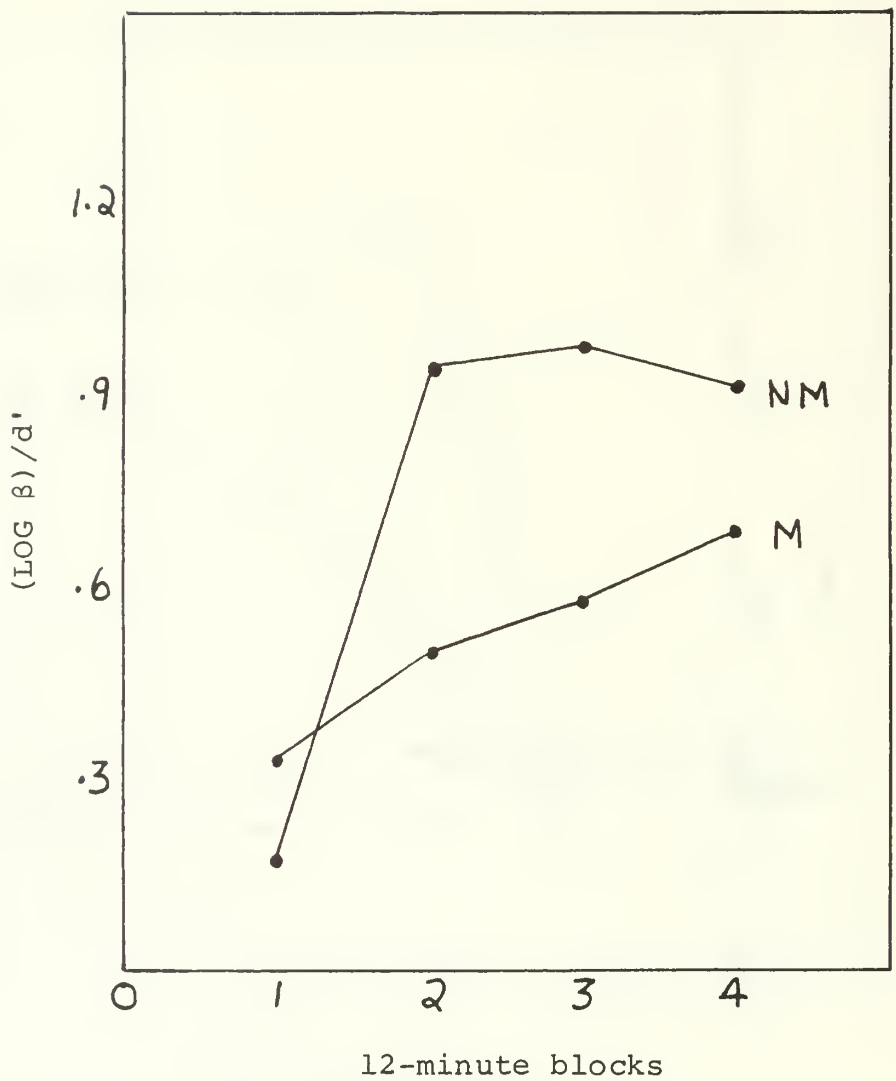

FIGURE 4: THE QUANTITY (LOG β)/d'
AS A FUNCTION OF TIME

BIBLIOGRAPHY

Baker, R. A., Jr., Kowal, B., and Ware, R. J., "The Role of Experimenter Attitude and Contingent Reinforcement in a Vigilance Task," Journal of the Human Factors Society, v. 6, p. 111-115, February 1966.

Bergum, B. O., and Lehr, D. J., "Monetary Incentives and Vigilance," Journal of Experimental Psychology, v. 67, p. 197-198, February 1964.

Deese, J., "Some Problems in the Theory of Vigilance," Psychology Review, v. 5, p. 359-368, September 1955.

Egan, J. P., Greenberg, G. Z., and Schulman, A. I., "Operating Characteristics, Signal Detectability, and the Method of Free Response," Journal of Acoustical Society of America, v. 33, p. 933-1007, July 1961.

Galanter, E., "The Direct Measurement of Utility and Subjective Probability," American Journal of Psychology, v. 75, p. 208-220, June 1962.

Green, D. M., and Swets, J. A., Signal Detection Theory and Psychophysics, Wiley, 1966.

Hicks, C. R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart, and Winston, 1964.

Jerison, H. J., and Pickett, R. M., "Vigilance: A Review and Re-evaluation," Journal of the Human Factors Society, v. 5, p. 211-238, June 1963.

Jerison, H. J., Pickett, R. M., and Stenson, H. H., "The Elicited Observing Rate and Decision Processes in Vigilance," Journal of the Human Factors Society, v. 7, p. 107-128, April 1965.

Levine, J. M., "The Effects of Values and Costs on the Detection and Identification of Signals in Auditory Vigilance," Journal of the Human Factors Society, v. 8, p. 525-537, February 1966.

Loeb, M. and Binford, J. R., "Vigilance for Auditory Intensity Changes as a Function of Preliminary Feedback and Confidence Level," Journal of the Human Factors Society, v. 6, p. 445-458, October 1964.

Mackworth, J., "The Effect of Signal Rate on Performance in Two Kinds of Vigilance Task," Journal of the Human Factors Society, v. 10, p. 11-17, February 1968.

Pollack, I., and Knaff, P. R., "Maintainance of Alertness by a Loud Auditory Signal," Journal of the Acoustical Society of America, v. 30, p. 1013-1016, November 1958.

Selin, I., Detection Theory, Rand Corporation, 1965.

Seigel, S., Nonparametric Statistics, McGraw-Hill, 1956.

Sipowicz, R. R., Ware, J. R., and Baker, R. A., "The Effects of Reward and Knowledge of Results on Performance of a Single Vigilance Task," Journal of Experimental Psychology, v. 64, p. 58-61, January 1962.

Smith, R. L., Lucaccini, L. F., and Epstein, M. H., "Effects of Monetary Rewards and Punishment on Vigilance Performance," Journal of Applied Psychology, v. 51, p. 411-416, December 1967.

Swets, J. A., Signal Detection and Recognition by Human Observers, Wiley, 1964.

Taylor, M. M., "Detectability Measures in Vigilance: Comment on a Paper by Wiener, Poock, and Steele, "Perception and Motor Skills, v. 20, p. 1217-1221, June 1965.

Wiener, E. L., "Knowledge of Results and Signal Rate in Monitoring: A Transfer of Training Approach," Journal of Applied Psychology, v. 47, p. 214-222, August 1963.

Wiener, E. L., Poock, G. K., and Steele, M., "Effect of Time-Sharing on Monitoring Performance: Simple Mental Arithmetic as a Loading Task," Perception and Motor Skills, v. 19, p. 435-440, October 1964.

Winer, B. J., Statistical Principles in Experimental Design, McGraw-Hill, 1962.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Chief of Naval Operations (OP 96) Department of the Navy Washington, D. C. 20350	1
4.	Asst. Professor G. K. Poock, Code 55Pk Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	3
5.	Assoc. Professor R. N. Forrest, Code 55Ft Department of Operations Analysis Naval Postgraduate School Monterey, California 93940	1
6.	Lt. Kenneth Lee Yufer, USN 211 James Street Flemington, Pennsylvania 17745	1

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

The Effect of Monetary Incentives on Monitoring Performance

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Master's Thesis; October, 1969

5. AUTHOR(S) (First name, middle initial, last name)

Kenneth Lee Yufer

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1969	30	22

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

A vigilance experiment was performed to evaluate the effect of monetary incentives on visual monitoring performance. The vigilance task was the detection of a slightly larger excursion of a voltmeter needle making 50 uniform excursions per minute. The length of the vigil was 48 minutes, during which 32 signals were presented. Ten subjects (Ss) in a control group performed the task without possibility of reward. A second group of 10 Ss performed the identical task receiving monetary rewards based on performance. Subjects receiving monetary rewards (M) detected significantly more signals ($p < .001$) than did the control group (NM). A significant time decrement ($p < .001$) and a significant interaction between group and time existed ($p < .001$). The detectability (d') or sensitivity for the signals remained essentially invariant for both groups, although slightly greater for group M, and the criterion level (beta) was comparatively lower for group M. The cost factors were effective in manipulating monitoring performance.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6811

Unclassified
Security Classification

A-31408

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Monitoring Performance						
Effect of Monetary Incentives						